Confusion Matrix Metrics Cheat Sheet

Confusion Matrix

A confusion matrix for a classifier is typically given by:

	Predicted: P	Predicted: N
Actual: P	TP	FN
Actual: N	FP	TN

- **True Positives (TP)**: The number of correct positive predictions.
- **False Positives (FP)**: The number of incorrect predictions where negative instances are classified as positive.
- **False Negatives (FN)**: The number of incorrect predictions where positive instances are classified as negative.
- **True Negatives (TN)**: The number of correct negative predictions.
- **Type I Error**: FP;
  Type II Error: FN

Confusion Matrix Example

Consider a dataset with two classes: Positive (P) and Negative (N). The confusion matrix is given by:

	Predicted: P	Predicted: N
Actual: P	50	10
Actual: N	5	100

Metrics

- **Sensitivity or Recall**: The ratio of TP to the actual positives:

$$\text{Sensitivity} = \frac{TP}{TP + FN}$$

- **Specificity**: The ratio of TN to the actual negatives:

$$\text{Specificity} = \frac{TN}{TN + FP}$$

- **Precision**: The ratio of TP to all predicted positives:

$$\text{Precision} = \frac{TP}{TP + FP}$$

- **Accuracy**: The ratio of correct predictions to the total number of cases:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

- **F1-Score**: The score calculating the balance between Precision and Recall. A good model will have a high F1-score:

$$\text{F1-Score} = 2 * \frac{Precision * Recall}{Precision + Recall}$$

Calculating Metrics from the Example

Given the confusion matrix above, we can calculate the following:

- Sensitivity (Recall): $\frac{50}{50+10} = 0.833$
- Specificity: $\frac{100}{100+5} = 0.952$
- Precision: $\frac{50}{50+5} = 0.909$
- Accuracy: $\frac{50+100}{50+10+5+100} = 0.909$
- F1-Score: $2 * \frac{0.909*0.833}{0.909+0.833} = 0.869$